

SEQUENCE LISTING.txt

SEQUENCE LISTING

<110> Omary, Bishr
Ku, Nam-Om

<120> Keratin 8 and 18 mutations are Risk
Factors for Developing Liver Disease of Multiple Etiologies

<130> STAN-297WO

<140> 10/552,949

<141> 2006-07-27

<150> 60/462,989

<151> 2003-04-14

<160> 9

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1485

<212> DNA

<213> H. sapiens

<220>

<221> CDS

<222> (119)...(1408)

<400> 1

tccggggcgg	gggcggggcc	tcactctgcg	atataactcg	ggtcgcgcg	ctcgcgcagg	60
ccgccaccgt	cgctcgcaaa	gcctgagtc	tgctctttct	ctctccccgg	acagcatg	118
agc ttc acc	act cgc tcc	acc ttc tcc	acc aac tac	cgg tcc ctg	ggc	166
Ser Phe Thr	Thr Arg Ser	Thr Phe Ser	Thr Asn Tyr	Arg Ser	Leu Gly	
1	5		10		15	

tct gtc	cag gcg	ccc agc	tac ggc	gcc cgg	ccg gtc	agc agc	gcg gcc	214
Ser Val	Gln Ala	Pro Ser	Tyr Gly	Ala Arg	Pro Val	Ser Ser	Ala Ala	
	20			25		30		

agc gtc	tat gca	ggc gct	ggg ggc	tct ggt	tcc cgg	atc tcc	gtg tcc	262
Ser Val	Tyr Ala	Gly Ala	Gly Gly	Ser Gly	Ser Arg	Ile Ser	Val Ser	
	35		40		45			

cgc tcc	acc agc	ttc agg	ggc ggc	atg ggg	tcc ggg	ggc ctg	gcc acc	310
Arg Ser	Thr Ser	Phe Arg	Gly Gly	Met Gly	Ser Ser	Gly Leu	Ala Thr	
	50		55		60			

ggg ata	gcc ggg	ggt ctg	gca gga	atg gga	ggc atc	cag aac	gag aag	358
Gly Ile	Ala Gly	Gly Leu	Ala Gly	Met Gly	Ile Ile	Gln Asn	Glu Lys	
65		70		75			80	

gag acc	atg caa	agc ctg	aac gac	cgc ctg	gcc tct	tac ctg	gac aga	406
Glu Thr	Met Gln	Ser Ser	Leu Asn	Asp Arg	Leu Ala	Ser Tyr	Leu Asp	
		85		90		95		

gtg agg	agc ctg	gag acc	gag aac	cgg agg	ctg gag	agc aaa	atc cgg	454
Val Arg	Ser Leu	Glu Thr	Glu Asn	Arg Arg	Leu Glu	Ser Lys	Ile Arg	
	100			105		110		

gag cac	ttg gag	aag aag	gga ccc	cag gtc	aga gac	tgg agc	cat tac	502

SEQUENCE LISTING.txt

Glu	His	Leu	Glu	Lys	Lys	Gly	Pro	Gln	Val	Arg	Asp	Trp	Ser	His	Tyr	
		115					120					125				
ttc	aag	atc	atc	gag	gac	ctg	agg	gct	cag	atc	ttc	gca	aat	act	gtg	550
Phe	Lys	Ile	Ile	Glu	Asp	Leu	Arg	Ala	Gln	Ile	Phe	Ala	Asn	Thr	Val	
	130					135					140					
gac	aat	gcc	cgc	atc	gtt	ctg	cag	att	gac	aat	gcc	cgt	ctt	gct	gct	598
Asp	Asn	Ala	Arg	Ile	Val	Leu	Gln	Ile	Asp	Asn	Ala	Arg	Leu	Ala	Ala	
145				150						155				160		
gat	gac	ttt	aga	gtc	aag	tat	gag	aca	gag	ctg	gcc	atg	cgc	cag	tct	646
Asp	Asp	Phe	Arg	Val	Lys	Tyr	Glu	Thr	Glu	Leu	Ala	Met	Arg	Gln	Ser	
				165					170					175		
gtg	gag	aac	gac	atc	cat	ggg	ctc	cgc	aag	gtc	att	gat	gac	acc	aat	694
Val	Glu	Asn	Asp	Ile	His	Gly	Leu	Arg	Lys	Val	Ile	Asp	Thr	Thr	Asn	
			180					185					190			
atc	aca	cga	ctg	cag	ctg	gag	aca	gag	atc	gag	gct	ctc	aag	gag	gag	742
Ile	Thr	Arg	Leu	Gln	Leu	Glu	Thr	Glu	Ile	Glu	Ala	Leu	Lys	Glu	Glu	
		195					200					205				
ctg	ctc	ttc	atg	aag	aag	aac	cac	gaa	gag	gaa	gta	aaa	ggc	cta	caa	790
Leu	Leu	Phe	Met	Lys	Lys	Asn	His	Glu	Glu	Glu	Val	Lys	Gly	Leu	Gln	
		210				215					220					
gcc	cag	att	gcc	agc	tct	ggg	ttg	acc	gtg	gag	gta	gat	gcc	ccc	aaa	838
Ala	Gln	Ile	Ala	Ser	Ser	Gly	Leu	Thr	Val	Val	Val	Asp	Ala	Pro	Lys	
225				230					235					240		
tct	cag	gac	ctc	gcc	aag	atc	atg	gca	gac	atc	cgg	gcc	caa	tat	gac	886
Ser	Gln	Asp	Leu	Ala	Lys	Ile	Met	Ala	Asp	Ile	Arg	Ala	Gln	Tyr	Asp	
				245					250					255		
gag	ctg	gct	cgg	aag	aac	cga	gag	gag	cta	gac	aag	tac	tgg	tct	cag	934
Glu	Leu	Ala	Arg	Lys	Asn	Arg	Glu	Glu	Leu	Asp	Lys	Tyr	Trp	Ser	Gln	
			260					265					270			
cag	att	gag	gag	agc	acc	aca	gtg	gtc	acc	aca	cag	tct	gct	gag	gtt	982
Gln	Ile	Glu	Glu	Ser	Thr	Thr	Val	Val	Thr	Thr	Gln	Ser	Ala	Glu	Val	
		275					280					285				
gga	gct	gct	gag	acg	acg	ctc	aca	gag	ctg	aga	cgt	aca	gtc	cag	tcc	1030
Gly	Ala	Ala	Glu	Thr	Thr	Leu	Thr	Glu	Leu	Arg	Arg	Thr	Val	Gln	Ser	
		290				295					300					
ttg	gag	atc	gac	ctg	gac	tcc	atg	aga	aat	ctg	aag	gcc	agc	ttg	gag	1078
Leu	Glu	Ile	Asp	Leu	Asp	Ser	Met	Arg	Asn	Leu	Lys	Ala	Ser	Leu	Glu	
305				310						315				320		
aac	agc	ctg	agg	gag	gtg	gag	gcc	cgc	tac	gcc	cta	cag	atg	gag	cag	1126
Asn	Ser	Leu	Arg	Glu	Val	Glu	Ala	Arg	Tyr	Ala	Leu	Gln	Met	Glu	Gln	
				325					330					335		
ctc	aac	ggg	atc	ctg	ctg	cac	ctt	gag	tca	gag	ctg	gca	cag	acc	cgg	1174
Leu	Asn	Gly	Ile	Leu	Leu	His	Leu	Glu	Ser	Glu	Leu	Ala	Gln	Thr	Arg	
			340					345				350				
gca	gag	gga	cag	cgc	cag	gcc	cag	gag	tat	gag	gcc	ctg	ctg	aac	atc	1222
Ala	Glu	Gly	Gln	Arg	Gln	Ala	Gln	Glu	Tyr	Glu	Ala	Leu	Leu	Asn	Ile	
		355				360						365				

SEQUENCE LISTING.txt

aag gtc aag ctg gag gct gag atc gcc acc tac cgc cgc ctg ctg gaa 1270
 Lys Val Lys Leu Glu Ala Glu Ile Ala Thr Tyr Arg Arg Leu Leu Glu
 370 375 380

gat ggc gag gac ttt aat ctt ggt gat gcc ttg gac agc agc aac tcc 1318
 Asp Gly Glu Asp Phe Asn Leu Gly Asp Ala Leu Asp Ser Ser Asn Ser
 385 390 395 400

atg caa acc atc caa aag acc acc acc cgc cgg ata gtg gat ggc aaa 1366
 Met Gln Thr Ile Gln Lys Thr Thr Thr Arg Arg Ile Val Asp Gly Lys
 405 410 415

gtg gtg tct gag acc aat gac acc aaa gtt ctg agg cat taa 1408
 Val Val Ser Glu Thr Asn Asp Thr Lys Val Leu Arg His *
 420 425

gccagcagaaa gcagggtacc ctttggggag caggaggcca ataaaaagt cagagtccaa 1468
 aaaaaaaaaa aaaaaaa 1485

<210> 2
 <211> 429
 <212> PRT
 <213> H. sapiens

<400> 2
 Ser Phe Thr Thr Arg Ser Thr Phe Ser Thr Asn Tyr Arg Ser Leu Gly
 1 5 10 15
 Ser Val Gln Ala Pro Ser Tyr Gly Ala Arg Pro Val Ser Ser Ala Ala
 20 25 30
 Ser Val Tyr Ala Gly Ala Gly Gly Ser Gly Ser Arg Ile Ser Val Ser
 35 40 45
 Arg Ser Thr Ser Phe Arg Gly Gly Met Gly Ser Gly Gly Leu Ala Thr
 50 55 60
 Gly Ile Ala Gly Gly Leu Ala Gly Met Gly Gly Ile Gln Asn Glu Lys
 65 70 75 80
 Glu Thr Met Gln Ser Leu Asn Asp Arg Leu Ala Ser Tyr Leu Asp Arg
 85 90 95
 Val Arg Ser Leu Glu Thr Glu Asn Arg Arg Leu Glu Ser Lys Ile Arg
 100 105 110
 Glu His Leu Glu Lys Lys Gly Pro Gln Val Arg Asp Trp Ser His Tyr
 115 120 125
 Phe Lys Ile Ile Glu Asp Leu Arg Ala Gln Ile Phe Ala Asn Thr Val
 130 135 140
 Asp Asn Ala Arg Ile Val Leu Gln Ile Asp Asn Ala Arg Leu Ala Ala
 145 150 155 160
 Asp Asp Phe Arg Val Lys Tyr Glu Thr Glu Leu Ala Met Arg Gln Ser
 165 170 175
 Val Glu Asn Asp Ile His Gly Leu Arg Lys Val Ile Asp Asp Thr Asn
 180 185 190
 Ile Thr Arg Leu Gln Leu Glu Thr Glu Ile Glu Ala Leu Lys Glu Glu
 195 200 205
 Leu Leu Phe Met Lys Lys Asn His Glu Glu Glu Val Lys Gly Leu Gln
 210 215 220
 Ala Gln Ile Ala Ser Ser Gly Leu Thr Val Glu Val Asp Ala Pro Lys
 225 230 235 240
 Ser Gln Asp Leu Ala Lys Ile Met Ala Asp Ile Arg Ala Gln Tyr Asp
 245 250 255
 Glu Leu Ala Arg Lys Asn Arg Glu Glu Leu Asp Lys Tyr Trp Ser Gln
 260 265 270
 Gln Ile Glu Glu Ser Thr Thr Val Val Thr Thr Gln Ser Ala Glu Val
 275 280 285
 Gly Ala Glu Thr Thr Leu Thr Glu Leu Arg Arg Thr Val Gln Ser
 290 295 300

SEQUENCE LISTING.txt

Leu Glu Ile Asp Leu Asp Ser Met Arg Asn Leu Lys Ala Ser Leu Glu
 305 310 315 320
 Asn Ser Leu Arg Glu Val Glu Ala Arg Tyr Ala Leu Gln Met Glu Gln
 325 330 335
 Leu Asn Gly Ile Leu Leu His Leu Glu Ser Glu Leu Ala Gln Thr Arg
 340 345 350
 Ala Glu Gly Gln Arg Gln Ala Gln Glu Tyr Glu Ala Leu Leu Asn Ile
 355 360 365
 Lys Val Lys Leu Glu Ala Glu Ile Ala Thr Tyr Arg Leu Leu Glu
 370 375 380
 Asp Gly Glu Asp Phe Asn Leu Gly Asp Ala Leu Asp Ser Ser Asn Ser
 385 390 395 400
 Met Gln Thr Ile Gln Lys Thr Thr Thr Arg Ile Val Asp Gly Lys
 405 410 415
 Val Val Ser Glu Thr Asn Asp Thr Lys Val Leu Arg His
 420 425

<210> 3
 <211> 1752
 <212> DNA
 <213> H. sapiens

<220>
 <221> CDS
 <222> (63)...(1511)

<400> 3
 ctgctccttc taggatctcc gccctggttcg gcccgccctgc ctccactcct gccctcacca 60
 tg tcc atc agg gtg acc cag aag tcc tac aag gtg tcc acc tct ggc 107
 Ser Ile Arg Val Thr Gln Lys Ser Tyr Lys Val Ser Thr Ser Gly 15
 1 5 10
 ccc cgg gcc ttc agc agc cgc tcc tac acg agt ggg ccc ggt tcc cgc 155
 Pro Arg Ala Phe Ser Ser Arg Ser Tyr Thr Ser Ser Gly Pro Gly Ser Arg 30
 20 25
 atc agc tcc tcg agc ttc tcc cga gtg ggc agc agc aac ttt cgc ggt 203
 Ile Ser Ser Ser Ser Phe Ser Arg Val Gly Ser Ser Asn Phe Arg Gly 40
 35 45
 ggc ctg ggc ggc ggc tat ggt ggg gcc agc ggc atg gga ggc atc acc 251
 Gly Leu Gly Gly Gly Tyr Gly Gly Ala Ser Gly Met Gly Gly Ile Thr 50
 55 60
 gca gtt acg gtc aac cag agc ctg ctg agc ccc ctt gtc ctg gag gtg 299
 Ala Val Thr Val Asn Gln Ser Leu Leu Ser Pro Leu Val Leu Glu Val 65
 70 75
 gac ccc aac atc cag gcc gtg cgc acc cag gag aag gag cag atc aag 347
 Asp Pro Asn Ile Gln Ala Val Arg Thr Gln Glu Lys Glu Gln Ile Lys 80
 85 90 95
 acc ctc aac aac aag ttt gcc tcc ttc ata gac aag gta cgg ttc ctg 395
 Thr Leu Asn Asn Lys Phe Ala Ser Phe Ile Asp Lys Val Arg Phe Leu 100
 105 110
 gag cag cag aac aag atg ctg gag acc aag tgg agc ctc ctg cag cag 443
 Glu Gln Gln Asn Lys Met Leu Glu Thr Lys Trp Ser Leu Leu Gln Gln 115
 120 125
 cag aag acg gct cga agc aac atg gac aac atg ttc gag agc tac atc 491
 Gln Lys Thr Ala Arg Ser Asn Met Asp Asn Met Phe Glu Ser Tyr Ile
 130 135 140 145

SEQUENCE LISTING.txt

130		135	140	
aac aac ctt agg cgg cag ctg gag act ctg ggc cag gag aag ctg aag	539			
Asn Asn Leu Arg Arg Gln Leu Glu Thr Leu Gly Gln Glu Lys Leu Lys				
145		150	155	
ctg gag gcg gag ctt ggc aac atg cag ggg ctg gtg gag gac ttc aag	587			
Leu Glu Ala Glu Leu Gln Asn Met Gln Gly Leu Val Glu Asp Phe Lys				
160		165	170	175
aac aag tat gag gat gag atc aat aag cgt aca gag atg gag aac gaa	635			
Asn Lys Tyr Glu Asp Glu Ile Asn Lys Arg Thr Glu Met Glu Asn Glu				
180		185	190	
ttt gtc ctc atc aag aag gat gtg gat gaa gct tac atg aac aag gta	683			
Phe Val Leu Ile Lys Lys Asp Val Asp Glu Ala Tyr Met Asn Lys Val				
195		200	205	
gag ctg gag tct cgc ctg gaa ggg ctg acc gac gag atc aac ttc ctc	731			
Glu Leu Glu Ser Arg Leu Glu Gly Leu Thr Asp Glu Ile Asn Phe Leu				
210		215	220	
agg cag cta tat gaa gag gag atc cgg gag ctg cag tcc cag atc tcg	779			
Arg Gln Leu Tyr Glu Glu Glu Ile Arg Glu Leu Gln Ser Gln Ile Ser				
225		230	235	
gac aca tct gtg gtg ctg tcc atg gac aac agc cgc tcc ctg gac atg	827			
Asp Thr Ser Val Val Leu Ser Met Asp Asn Ser Arg Ser Leu Asp Met				
240		245	250	255
gac agc atc att gct gag gtc aag gca cag tac gag gat att gcc aac	875			
Asp Ser Ile Ile Ala Glu Val Lys Ala Gln Tyr Glu Asp Ile Ala Asn				
260		265	270	
cgc agc cgg gct gag gct gag agc atg tac cag atc aag tat gag gag	923			
Arg Ser Arg Ala Glu Ala Glu Ser Met Tyr Gln Ile Lys Tyr Glu Glu				
275		280	285	
ctg cag agc ctg gct ggg aag cac ggg gat gac ctg cgg cgc aca aag	971			
Leu Gln Ser Leu Ala Gly Lys His Gly Asp Asp Leu Arg Arg Thr Lys				
290		295	300	
act gag atc tct gag atg aac cgg aac atc agc cgg ctc cag gct gag	1019			
Thr Glu Ile Ser Glu Met Asn Arg Asn Ile Ser Arg Leu Gln Ala Glu				
305		310	315	
att gag ggc ctc aaa ggc cag agg gct tcc ctg gag gcc gcc att gca	1067			
Ile Glu Gly Leu Lys Gln Arg Gln Arg Ala Ser Leu Glu Ala Ala Ile Ala				
320		325	330	335
gat gcc gag cag cgt gga gag ctg gcc att aag gat gcc aac gcc aag	1115			
Asp Ala Glu Gln Arg Gly Glu Leu Ala Ile Lys Asp Ala Asn Ala Lys				
340		345	350	
ttg tcc gag ctg gag gcc gcc ctg cag cgg gcc aag cag gac atg gcg	1163			
Leu Ser Glu Leu Glu Ala Ala Leu Gln Arg Ala Lys Gln Asp Met Ala				
355		360	365	
cgg cag ctg cgt gag tac cag gag ctg atg aac gtc aag ctg gcc ctg	1211			
Arg Gln Arg Glu Tyr Gln Glu Leu Met Asn Val Lys Leu Ala Leu				
370		375	380	
gac atc gag atc gcc acc tac agg aag ctg ctg gag ggc gag gag agc	1259			

SEQUENCE LISTING.txt

195 200 205
 Leu Glu Ser Arg Leu Glu Gly Leu Thr Asp Glu Ile Asn Phe Leu Arg
 210 215 220
 Gln Leu Tyr Glu Glu Ile Arg Glu Leu Gln Ser Gln Ile Ser Asp
 225 230 235
 Thr Ser Val Val Leu Ser Met Asp Asn Ser Arg Ser Leu Asp Met Asp
 240 245 250
 Ser Ile Ile Ala Glu Val Lys Ala Gln Tyr Glu Asp Ile Ala Asn Arg
 255 260 265
 Ser Arg Ala Glu Ala Glu Ser Met Tyr Gln Ile Lys Tyr Glu Glu Leu
 270 275 280
 Gln Ser Leu Ala Gly Lys His Gly Asp Asp Leu Arg Arg Thr Lys Thr
 285 290 300
 Glu Ile Ser Glu Met Asn Arg Asn Ile Ser Arg Leu Gln Ala Glu Ile
 305 310 315
 Glu Gly Leu Lys Gly Gln Arg Ala Ser Leu Glu Ala Ala Ile Ala Asp
 320 325 330
 Ala Glu Gln Arg Gly Glu Leu Ala Ile Lys Asp Ala Asn Ala Lys Leu
 335 340 345
 Ser Glu Leu Glu Ala Ala Leu Gln Arg Ala Lys Gln Asp Met Ala Arg
 350 355 360
 Gln Leu Arg Glu Tyr Gln Glu Leu Met Asn Val Lys Leu Ala Leu Asp
 365 370 375
 Ile Glu Ile Ala Thr Tyr Arg Lys Leu Leu Glu Gly Glu Glu Ser Arg
 380 385 390
 Leu Glu Ser Gly Met Gln Asn Met Ser Ile His Thr Lys Thr Thr Ser
 395 400 405
 Gly Tyr Ala Gly Gly Leu Ser Ser Ala Tyr Gly Gly Leu Thr Ser Pro
 410 415 420
 Gly Leu Ser Tyr Ser Leu Gly Ser Ser Phe Gly Ser Gly Ala Gly Ser
 425 430 435
 Ser Ser Phe Ser Arg Thr Ser Ser Ser Arg Ala Val Val Val Lys Lys
 440 445 450
 Ile Glu Thr Arg Asp Gly Lys Leu Val Ser Glu Ser Ser Asp Val Leu
 455 460 465
 Pro Lys 470 475 480

<210> 5
 <211> 26
 <212> DNA
 <213> H. sapiens

<400> 5
 atgccgagca gcgtggagag ctggcc

26

<210> 6
 <211> 7
 <212> PRT
 <213> H. sapiens

<400> 6
 Ala Glu Gln Arg Gly Glu Leu
 1 5

<210> 7
 <211> 26
 <212> DNA
 <213> H. sapiens

<220>

SEQUENCE LISTING.txt

<221> mutation

<222> 13

<223> N = any nucleotide

<400> 7

atgccgagca gcntggagag ctggcc

26

<210> 8

<211> 7

<212> PRT

<213> H. sapiens

<220>

<221> VARIANT

<222> 4

<223> Xaa = R or H

<400> 8

Ala Glu Gln Xaa Gly Glu Leu

1

5

<210> 9

<211> 14

<212> PRT

<213> H. sapiens

<400> 9

Ile Ala Asp Ala Glu Gln Arg Gly Glu Leu Ala Ile Lys Asp

1

5

10